DOCUMENT RESUME

ED 414 884 IR 018 654

AUTHOR Miller, David W.; Lekander, Brian
TITLE Extending and Disseminating a New Method of Teaching U.S. History.
INSTITUTION Carnegie-Mellon Univ., Pittsburgh, PA. Dept. of History.
SPONS AGENCY Fund for the Improvement of Postsecondary Education (ED), Washington, DC.
PUB DATE 1992-00-00
NOTE 24p.
CONTRACT P116B00229
PUB TYPE Reports - Descriptive (141)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Access to Information; *Census Figures; *Computer Software; Computer Uses in Education; Demography; Elections; Higher Education; History Instruction; Hypermedia; Information Dissemination; *Online Searching; Profiles; *Statistical Data; Teaching Methods; Undergraduate Students; *United States History

ABSTRACT

This report describes a project intended to overcome certain obstacles to the dissemination to other postsecondary institutions of a new technique for teaching U.S. History developed at Carnegie Mellon University. This technique utilizes a software system, The Great American History Machine (GAHM), which enables undergraduates to explore a huge body of census and elections data for the counties of the United States for the 19th and 20th centuries through a map interface. The purpose of this teaching technique is to empower undergraduates to think like historians. Professional historians search for patterns in huge bodies of information which they master in the course of years of studying a particular society or problem in the past. Teacher documentation was written, new datasets were prepared for inclusion in the system, and a hypertext capability was developed to enable students to search the system in more intuitive ways and to consult online information on the meaning of census variables included in the system. Arrangements have been made for the software to be maintained and distributed by the Academic Software Development Group of the University of Maryland. An executive summary is provided. (SWC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

HE

Cover Sheet

Grantee Organization:

Carnegie Mellon University
Department of History
Schenley Park
Pittsburgh, PA 15213

ED 414 884

Grant Number:

P116B91025

Project Dates:

Starting Date: August 15, 1989
Ending Date: February 15, 1992
Number of Months: 30

Project Director:

David W. Miller
Department of History
Carnegie Mellon University
Schenley Park
Pittsburgh, PA 15213
Telephone: (412) 268-2880

FIPSE Program Officer: Brian Lekander

Grant Award:

Year 1	$ 50,613
Year 2	51,390
Cost extension	5,032
Total	$107,035

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

IR018654

Extending and Disseminating a New Method of Teaching U.S. History

Prior to the grant period we had invented a teaching technique using the Great American History Machine, software developed at CMU which enables students to access huge amounts of social, economic and political data for the nineteenth and twentieth centuries through a map interface. FIPSE funding enabled us to enhance this product in ways designed to facilitate its use outside CMU. Additional datasets have been prepared, teachers' documentation written, and an on-line help system to access text data developed. Finished products should be released shortly by the Academic Software Development Group of the University of Maryland.

Project director:

David W. Miller
Department of History
Carnegie Mellon University
Pittsburgh, PA 15213

(412) 268-2880

Publications and products:

David W. Miller, "Visualizing Patterns in Data: The Great American History Machine," forthcoming in a book on educational computing in the humanities being edited by Robert B. Kozma and Jerome Johnston of the National Center for Research to Improve Postsecondary Teaching and Learning.

David W. Miller, et al, the Great American History Machine, enhanced versions for Windows and for Macintosh. (Title subject to change on basis of marketing considerations.) Forthcoming from the Academic Software Development Group of the University of Maryland.

Executive Summary

Project title:

Extending and Disseminating a New Method of Teaching U.S. History

Grantee organization:

Carnegie Mellon University
Department of History
Schenley Park
Pittsburgh, PA 15213

Project director:

David W. Miller
(412) 268-2880

A. Project Overview.

The project was intended to overcome certain obstacles to the dissemination to other postsecondary institutions of a new technique for teaching U.S. History developed at Carnegie Mellon University by Project Director Miller and his colleague Professor John Modell. This technique utilizes a software system, The Great American History Machine (GAHM), developed by Miller which enables undergraduates to explore a huge body of census and elections data for the counties of the United States for the nineteenth and twentieth centuries through a map interface.

Teacher documentation was written, new datasets were prepared for inclusion in the system and a hypertext capability was developed to enable students to search the system in more intuitive ways and to consult online information on the meaning of census variables included in the system. Overall success in this effort is evidenced by the fact that arrangements have been made for the software to be maintained and distributed by the Academic Software Development Group of the University of Maryland.

B. Purpose.

The purpose of the teaching technique which this project seeks to disseminate is to empower undergraduates to think like historians. Professional historians, when they are practicing their craft, search for patterns in huge bodies of information which they master in the course of years of studying a particular society or problem in the past . Obviously it is not easy to apply this model to student learning in the course of a semester.

Our innovation seeks to simulate the professional historian's activity in the experience of an undergraduate by confronting that student with a truly massive body of information -- the county-level U.S. census results from 1790 to the present and the county-level election returns since 1840 -- through the medium of a software tool which facilitates the search for geographic patterns in those data.

C. Background and Origins.

The effort of which this project is a part began in 1985 when Miller was allocated an advanced function workstation as part of a CMU inititiative to encourage faculty to develop educational software on Andrew, the Unix-based distributed campus computing system being developed with IBM support. After Miller had, with some technical assistance, created a very crude prototype of GAHM, he persuaded CMU to allocate resources to employ a professional programmer to implement GAHM properly.

During the succeeding several years the software matured with both continued internal CMU support and funding from IBM raised by Miller. In 1987, Miller and Modell offered for the first time a sophomore course, "Pattern and Change in Nineteenth Century America" which is structured around a sequence of exercises in which students explore problems in U. S. history using GAHM. Our first preliminary FIPSE proposal, submitted in the fall of 1987, was rejected, but in the following year we were successful in obtaining FIPSE funding for the project of which this is the final report.

D. Project Description.

Two graduate students, working under Modell's direction, wrote a manual for teachers which contains six examples of classic problems in U.S. history which can fruitfully be explored by undergraduates

using GAHM. Our programmer created a new hypertext capability which facilitates access to the data in a more intuitive way by users not familiar with the history of the census and the topics on which it reports. This capability also allows users to access files which define and explain particular census variables and treat larger issues in a hierarchy of topics into which the variables have been organized. A team of students carried out the research and wrote such files for many problematic variables. Additional datasets and the maps to accomodate them were prepared for inclusion in the system.

In the body of the report we give details on both successes and difficulties in the execution of these tasks. Important issues include the relationship of this effort to changing goals and strategies at CMU with respect to educational computing, and the tension between intensive educational development and the discipline-based research in the role of a faculty member at a research-oriented university.

E. Project Results.

The most important result was that we have succeeded in making an arrangement for the maintenance, upgrading and dissemination of GAHM, as enhanced with FIPSE support, by the Academic Software Development Group of the University of Maryland. This should ensure the attainment of the primary goal of the funding, the actual dissemination of our teaching technique to postsecondary institutions other than CMU. Since this was a technology development effort rather than a classroom-based experiment, we cannot yet report specifics on its impact..

F. Summary and Conclusions.

We have significantly improved the software on which our new method for teaching U.S. history is based, and we have taken steps which should make it genuinely accessible in a wide variety of postsecondary institutions.

G. Appendices.

In an appendix we offer some suggestions to FIPSE and potential FIPSE grantees about budgeting for programming efforts and about the special problems which arise when the type of effort FIPSE supports involves technology development by subject-matter specialists.

Final Report on FIPSE Project P116B91025 "Extending and Disseminating a New Method for Teaching U.S. History"

A. Project Overview.

This project grew out of an instructional innovation developed at CMU during 1987-89 by David W. Miller and John Modell. These two members of the History Department created a sophomore course in nineteenth century U.S. history which uses a software system, the Great American History Machine (GAHM), which Miller had been developing since 1985. GAHM enables students to explore historical U.S. census data through a map interface, and the course, "Pattern and Change in Nineteenth Century America" is built around a sequence of exercises which invite students to formulate hypotheses based on such exploration.

The FIPSE grant was intended to overcome some of the obstacles to the adoption of this teaching technique at postsecondary institutions other than CMU. Such obstacles included

1. the fact that GAHM was available only on certain Unix workstations favored by computer scientists rather than the more popular DOS and Macintosh computers typically available for student use at most relevant institutions.

2. the fact that undergraduates at most institutions may not have as much instruction in computing and access to as much formal and informal help in computer use as is the case at CMU.

3. the lack of documentation which would help faculty to envisage the kinds of assignments for which the software can be used in U.S. history courses.

4. the lack of easily accessible information on the structure and meaning of the data (which was mitigated at CMU by Miller

and Modell's detailed knowledge of the history of the census -- knowledge which many able U.S. history instructors would lack).

5. a few limitations in the availability of data, notably the greater emphasis on census rather than election data in the choice of datasets because of the CMU history department's focus on social rather than political history.

The project was designed specifically to address these obstacles. FIPSE funding was intended primarily to address items 2 - 5, and CMU gave assurances that item 1 would be addressed through internal support for a port of GAHM to the Macintosh. For the most part the original plans have been carried out, as will be indicated in detail below.

This project is perhaps unusual among FIPSE projects in that the audience it seeks primarily to serve is not located at the institution in which the work is being carried out. While much of the work done with FIPSE support will be incorporated in the local implementation of GAHM in CMU teaching situations, actual instruction at CMU has continued to use the stable version of the software which existed when the project began. Thus, in a sense, the audience for the project is invisible to us, or rather visible mainly in the form of a large file of requests to be informed when the software becomes available on Macintosh or DOS platforms.

A very important outcome of our efforts during the period of the grant is an evolving relationship with the Academic Software Development Group, headed by Dr. Chad McDaniel, at the University of Maryland. The relationship originated in discussions between Miller and McDaniel over difficulties which had been encountered in obtaining support to port the software to DOS environments at CMU. The fact that Maryland had substantial IBM funding for DOS-based educational software development (whereas it seemed impossible to obtain IBM funding at CMU for non-Unix development) led to a decision to conduct the DOS port at Maryland.

The Maryland group has now completed a Windows version of GAHM and arrangements have been made for them to handle the maintenance, upgrading and distribution of the software on both Windows and Macintosh platforms and of the accompanying documentation. We are especially pleased that Stephan Greene, the

programmer of both the Unix and Macintosh versions, will be joining the Maryland group at the end of the summer of 1992. The arrangement with the Maryland group greatly improves the prospects that the work which FIPSE has supported will actually have its intended impact beyond CMU.

B. Purpose

The fundamental purpose of the teaching technique which this project seeks to disseminate is to empower undergraduates to think like historians. What the historian primarily does, we believe, is to search for patterns in massive amounts of evidence. Unlike other social scientists, who identify hypotheses and devote their main professional efforts to operationalizing them and subjecting them to formal tests, the historian typically devotes his or her main professional attention to *generating* hypotheses.

In practice what this means is that the working historian may spend decades building a huge mental "database," for example by reading everything available which was written during the first half of the seventeenth century in France, so that he or she can discern patterns and formulate hypotheses which could not be found by anyone who had not made this enormous investment in mastering the evidence. (This, indeed, is why historians are so much more reluctant that other scholars to switch from one field to another.)

It might seem hopeless to try to give undergraduates, within the confines of a semester, the kind of experience described above. Certainly one common technique for getting students to think "like historians" -- the use of a collection of primary documents alongside a textbook -- fails precisely because the encounter with pre-selected pieces of evidence is so unlike what the professional historian's confrontation with raw evidence. Rather than generating new hypotheses by searching for patterns in the evidence the student usually does no more than find examples to corroborate what he or she has read in the textbook.

The paperback revolution of the 1950s was used by postsecondary teachers of history as an alternative to the textbook-plus-documents model for enabling students to think like historians. Rather than try to simulate the historian's main task within the experience of the student, teachers of history, at least in elite

institutions, tried to give the student the experience of looking over the shoulder of a master historian at work. In other words, students would read excellent, usually recent, research monographs by outstanding historians, and class discussion would be devoted largely to how the historian came to know what he or she reports in the book.

Our teaching technique resurrects the objective of simulating within the student's experience the search for patterns in evidence, but it substitutes for the compact collection of pre-selected documents a truly massive body of evidence, namely US census data from 1840 to 1980 (plus, thanks to FIPSE support, the census data for 1790 to 1830 and election data since 1840) -- over 6000 variables -- for each of the counties of the United States (about 3000 counties by 1990). Thus the student has access to perhaps 15,000,000 pieces of data -- an ample amount of evidence to simulate the experience of the mature professional historian confronting his or her personal mental database.

The specific vehicle which the student uses for approaching these data is the choropleth map -- in particular a map of the U.S. in which the counties are shaded according to the values of a social, economic or political variable. New variables can be created by arithmetic manipulation of the archived variables, class intervals for different shading patterns can be adjusted, maps can be zoomed and panned, and a map of a second variable, or of contemporary physical features, can be displayed alongside a given map. Even historians who are not temperamentally inclined to quantitative analysis often find this a congenial environment in which to explore data, and we believe that this is the case because it simulates the interactive search for patterns which is the historian's normal professional mode of problem-solving.

We were convinced before we obtained the FIPSE grant that we had largely achieved this purpose with our own students, although the inclusion of an online help system giving access especially to definitions of census variables, which does fall within the FIPSE project, was clearly desirable in order better to attain our original objectives. The FIPSE project had the more specific purpose of extending to postsecondary institutions other than CMU the techniques which we had developed to achieve the purpose outlined above.

Both our original pedagogical purpose and the more specific dissemination objectives which were the occasion for the FIPSE grant were well-specified at the beginning of the funding period. Circumstances have forced us to set some priorities with respect to aspects of the project, so that some detailed objectives have been more fully achieved than others, as will be discussed below. However, the overall purposes of the larger project and of its FIPSE-funded component have not changed significantly since the beginning of the funding period. No doubt this stability of purpose reflects the fact that the FIPSE project mainly entailed programming, data preparation, research and writing tasks, rather than classroom activity. Prior to FIPSE funding, as we were learning how to use the software with students, our understanding of our purpose evolved significantly; the description of that purpose which is set forth at the beginning of this section reflects the culmination, rather than the beginning, of that evolutionary process.

C. Background and Origins.

The effort of which this project is a part began in 1985 when Carnegie Mellon's Andrew project, which was funded by IBM, to develop a distributed campus computing system based on Unix workstations, had reached a point at which faculty could be asked to participate. Faculty were urged to submit proposals for educational software projects which might be implemented on advanced function workstations.

Miller, who had initiated several, less ambitious, educational computing efforts over the previous seven years, submitted such a proposal for what he called a "Great American History Machine." That proposal envisaged educational software for use in introductory U.S. history courses which would take advantage of three features of the proposed system: (1) very fast processors which permitted a high level of interactivity, (2) high-resolution monitors and (3) the storage and retrieval capabilities of the distributed file system. GAHM, as initially envisaged, would give students access to essentially all the available county-level U.S. census and election data from 1790 to the present through an interactive choropleth map interface.

In due course the proposal was accepted and Miller was supplied with a Sun 2 workstation (later replaced by an IBM RT when that machine was released), and was given a good deal of technical assistance particularly by Tom Peters, a senior IBM systems designer

then employed at the Information Technology Center which was developing the Andrew system. During 1985, with considerable technical assistance from Peters, Miller used his modest but non-trivial programming skills to create a very crude prototype of GAHM. The objective was to persuade CMU to allocate resources to employ a professional programmer to implement the project properly.

That objective was attained in the fall of 1985 when CMU's Center for the Design of Educational Computing assigned Stephan Greene, a recent CMU graduate, to the project part-time. Greene has been associated with the project from that time until 31 March, 1992, and he will resume a role in it in the late summer of 1992 when he joins the staff of the Academic Software Development Group at the University of Maryland. Throughout much of this period, however, he has devoted only part time to GAHM; one of the strategies for keeping the project going has been to find other projects, more or less closely related to GAHM, which could supplement the income stream which it was possible to raise for the GAHM project. Some consequences of this strategy will be discussed below.

By the fall of 1986 a working prototype of GAHM, containing the census variables for 1850 and a map of the U.S. by counties for 1850, had been implemented. About that time Miller and Modell began designing a sophomore-level course, "Pattern and Change in Nineteenth Century America," which would be structured around a series of GAHM-based exercises. The course was first taught in the fall semester of 1987, and during that semester we submitted a proposal to FIPSE requesting funding to support efforts to disseminate our teaching method beyond CMU; this proposal was rejected.

Meanwhile, from July 1986 to December 1987 a good deal of Miller's time was spent in efforts to secure direct funding from IBM (as opposed to indirect support through IBM's sponsorship of the Andrew project) for continued development of GAHM. Prominent among these efforts was participation in a number of IBM-sponsored demonstrations of GAHM at various sites. Eventually, in December 1987, IBM did commit $53,982 to the project.

IBM's principal reason for supporting GAHM at this point was to create a demonstratable application of the Andrew Tool Kit (ATK), an object-oriented programming environment whose development at CMU IBM had funded. Accordingly, much of the funding supplied by

IBM had to be used to re-implement the GAHM user interface in this new environment. This reprogramming effort certainly had some advantages; object-oriented programming does seem to be the wave of the future, and in the Macintosh version we have been able to retain much of the ATK-based program structure, while relying on a different object-oriented programming environment, MacApp. ATK itself, however, did not turn out to be the wave of the future, and the ATK-based code had to be almost completely rewritten by the Maryland group to make GAHM run on DOS-based IBM hardware.

The IBM funding did enable us to complete the preparation of the map with county boundary changes at decadal intervals from 1840 to 1980 and of census datasets for the same period. The user interface matured and the system became stable so that in the CMU context it could be relied upon to support our teaching application. During the 1988-89 academic year we submitted to FIPSE a second proposal for steps to facilitate effective dissemination of our educational innovations, and this time the proposal was funded.

D. Project Description:

This section is organized around the major tasks which were undertaken during the period of FIPSE funding.

D.1. Development of paper documentation

Perhaps the easiest component of the project to describe is the development of a teacher's manual. This task was undertaken by two CMU history graduate students, Timothy Haggerty and Joseph Spillane under the direction of John Modell. This group identified six themes and wrote a manual entitled *Student Research with the Great American History Machine: A Brief Manual with Six Good Examples.* The themes are:

-> The Turner Thesis in the American Imagination.

-> The Civil War and the Fall of the Second Party System.

-> Land and Labor in the New South.

-> The "Great Migration" of African-Americans from the Rural South to Northern Cities.

-> Critical Election Theory and the New Deal Coalition.

-> The Suburban World.

The themes were chosen to include both nineteenth and twentieth century material and to represent both social and political history. In each case the manual reviews scholarship in the area and poses questions for investigation with hints as to good strategies for getting students to explore those questions with GAHM. Suggestions for further reading are included for each theme.

FIPSE funding was used to support work on the manual by the two students for the summer of 1990 and part of the fall 1990 semester and to pay one month of Modell's salary. The manual was completed in November 1990, and a copy of it is submitted with this report. It is essentially complete, though the fact that it was written before the completion of the new user interface means that it may be necessary to do some very minor editing before releasing it. The manual has been supplied to the Maryland group, and they plan to incorporate it in their release of Windows and Macintosh versions of GAHM.

D.2. Programming tasks

There were essentially two programming objectives. The first was to create a hypertext capability which would offer friendlier approaches to the data than were available in the stable 1988 version, supply online textual information about the data and provide the infrastructure for a help system. The second was to port the software to the Macintosh platform. While these objectives are logically separable, we treat them here together because they became intertwined in practice.

The fact that Miller, despite being an historian by profession, already had more than four years experience in directing software development undoubtedly paid off in an understanding of how to set priorities in further development. The most important such lesson was to give priority to building a robust and versatile infrastructure rather than to a slick user interface. It is always easy to improve the user interface if the internals of the system are well designed. (Unfortunately, as will become clear, these lessons were not as well understood by others on whom we had some dependencies.)

Early in the period of FIPSE funding we realized that to build the hypertext capability we would need a library of subroutines for handling database management functions. We reviewed options and decided to use dbFile/dbVista products supplied by Raima Corporation. This development environment is available for DOS, Macintosh and Unix platforms. We have used it on our primary development platform, Unix, and in our port to the Macintosh. As the Maryland group incorporates the hypertext capabilities into the Windows version they will be able to use the DOS version of dbFile/dbVista to take advantage of our development work on the other platforms.

Much of the programming activity in the Unix environment through which the hypertext capability was developed took place between early 1990 and mid-1991. The result is a well-structured, modular, object-oriented, robust system with greatly extended capabilities for improving the user's access to the data and information about the data and about the user interface. These capabilities will be described more fully below under the heading "Preparation of online information."

The FIPSE project budget provided only part of the salary of the programmer, Stephan Greene, during each funding year. Accordingly, it was necessary to find other sources of funding to keep him in the university's employment throughout the project period -- a practice which was already well-established prior to receipt of the FIPSE grant, as we had seldom been able to fund him to work full-time on GAHM. During the first year of FIPSE funding the Information Technology Center (ITC) continued to employ Greene part-time to work on development and maintenance of the Andrew Tool Kit. This arrangement did have some payoffs for GAHM, as it enabled Greene sometimes to develop ATK functionalities which were needed in GAHM, but fundamentally it was, from our point of view, just a way to keep Greene fully employed in the absence of funding to keep him working full-time on GAHM.

In the early fall of 1990 we ended Greene's longstanding part-time relationship with the ITC because of the prospect of two other funding sources whose objectives seemed more central to those of the GAHM project. One of these sources was a grant from Sun Microsystems to Todd Cavalier, a faculty member in the Design Department in CMU's College of Fine Arts. Cavalier is using GAHM as

a test bed in which to do research and development on user interfaces, with the goal of creating an environment for scholarly collaboration on very advanced workstations. Essentially Greene worked full time for Cavalier for the first six months of 1991, though there were spinoff benefits for GAHM, because Cavalier's requirements included further work on the internals of the hypertext capability which was also essential to the FIPSE project.

Another new, though problematic, source of support which prompted our severing of Greene's relationship to the ITC in 1990 was the "Mac II Initiative" of CMU's College of Humanities and Social Sciences (H&SS). To explain how this matter arose we must recall some circumstances surrounding FIPSE's approval of the grant in 1989. In discussions leading to the final proposal, FIPSE had made it plain that it did not want to fund the port of the software to new platforms. However, FIPSE also insisted upon assurances that the software would be ported to at least one popular platform (DOS or Mac) with non-FIPSE support.

These discussions came at a point at which CMU was going through an important change. The decision to build the Andrew system had been accompanied by a determination by the central administration to achieve a leadership role in educational computing. Support for the initial development of GAHM had been a result of that initiative. In 1987 and 1988, however, Andrew was being perceived as having failed to achieve the goals of its most outspoken advocates, and, apart from GAHM, there were few examples of successful outcomes of the centralized sponsorship of innovation in educational computing. Educational computing efforts were radically decentralized (though the lessons of the centralized experiment were not always learned by those who bore responsibility for the decentralized effort). A major equipment grant from Apple was directed mainly toward H&SS, and that college was expected to bear the major responsibility for satisfying Apple's desire to see new educational applications result from this grant.

By the time of the discussions with FIPSE, GAHM had already been included in a list of projects to be supported by this "Mac II Initiative." The then Dean of H&SS, Stephen Fienberg, gave assurances in a letter to Miller (included as an appendix in our proposal) that "if your FIPSE proposal is funded, the College will give sufficient priority to this port of GAHM to ensure that it is completed no later than the end of the funding period of the FIPSE grant." At

an early stage of the Mac II Initiative Miller had tried to persuade the College to employ Greene half-time in that initiative -- an arrangement which seemed sensible in light of the Dean's stated strategy of employing enough programming staff to devote half of a programmer's time to each department's needs. The College took the position in this case (and in the case of at least one other department which had a closely analogous situation) that the Mac II initiative ought to be carried out only by full-time employees of the College.

The College's approach to the problem suffered from several flaws. The programmers employed for the Mac II initiative lacked experience on projects as ambitious as GAHM which required skill in advanced, object-oriented programming techniques, as opposed to fairly simple applications of packaged programming environments (such as Hypercard) which were appropriate for typical one-classroom faculty-generated instructional ideas. Moreover, the management of the initiative was often driven by deadlines for the production of demonstratable applications to visiting Apple representatives. This led to fragmentation of programmers' efforts and to a focus on making the user interface seem to work, to the neglect of the internals of the program. Moreover, the initiative suffered from some unfortunate turnover in its programming staff.

The result was that in the summer of 1990, Dean Fienberg had to admit that the effort to carry out the port of GAHM to the Macintosh with College programming staff had failed. Approximately one month's work by a College programmer who did have the requisite skills turned out to be salvageable from what was clearly an effort on which a number of person-months had been wasted. Dean Fienberg offered us two months of support for Greene to complete the task, an offer which Miller felt he had no choice but to accept, although he indicated to Fienberg and to his FIPSE Project Officer that this might not be adequate to finish the task.

A further difficulty was that we had not devoted sufficient thought at the time of our original proposal to the problem of incorporating into the Macintosh version the new hypertext capabilities which Greene was developing on Unix. Fienberg's position was that he was responsible only for porting GAHM as it had existed prior to 1989 to the Macintosh. Whether this was a fair reading of his letter of assurance is perhaps moot since he was offering a good deal less than needed to carry out even his restricted reading of that commitment.

It being obvious that a good deal more than two months of Greene's time would be required for Macintosh programming, Miller cobbled together a strategy to gain more of Greene's time for that purpose. A small amount of unexpended money was found in an earlier grant; an additional $5000 was raised internally from the CMU Division of Academic Services; a supplementary grant from FIPSE was requested and obtained; and Greene was directed to do as much development as possible directly on the Mac rather than on Unix. The utility of this last tactic was constrained by the fact that we needed to take as much advantage as possible of relevant work Greene might be able to do during the six months funded by Sun Microsystems, which of course had to be done in the Unix environment.

One further complication was that we had developed a relationship with SSD, the national social science data archive of Sweden, which contracted for some of Greene's time, including the period 1 November 1991 to 31 January 1992 which he spend in Gothenberg carrying out some programming tasks irrelevant to GAHM. Miller felt that it was necessary to cooperate in enabling Greene to take advantage of this opportunity for work abroad in order to avert the possible catastrophe of Greene leaving CMU before bringing the programming tasks to sufficient completion to enable us to turn the products over to a distributor (as matters worked out, to the Maryland group).

Despite all these complications the two major programming objectives noted at the beginning of this subsection -- the hypertext capability and the port of GAHM with that capability to the Macintosh -- were met by 31 March 1992 when Greene left the employment of CMU. It will be noted that this date is more than a month after the end of FIPSE funding. Non-FIPSE funds were of course used to pay Greene's compensation after 15 February, and those funds were essentially the last funds which we expect to raise for this effort. We are of course pleased to be able to report the attainment of the programming objectives. As will become evident, however, the overall objectives of the entire project would have been a bit better served if these specific programming tasks had been completed some months before the expiration of all funding for wrapping up CMU's role in the project.

D.3. Preparation of online information

A major obstacle to dissemination of our teaching technique beyond CMU was the fact that many U.S. history instructors, and certainly most students, would not have the level of knowledge about the U.S. census that Modell and Miller possessed by the time they began using GAHM for instruction. The start-up screen of the program as it existed prior to FIPSE funding invited the user to begin by choosing a dataset from a scrollable list: "1840 Census, 1844 Election . . . 1980 Census . . . ." If the user was interested in, say, immigration, the only way to find out which datasets contained data on that topic was to begin selecting one at a time and scrolling through the list of topics under which variables from that census had been, somewhat arbitrarily, grouped. A student who was interested in, say, social class, would have to figure out for him or herself that relevant information might be found under headings like "Housing."

A further difficulty was that the names of many variables were not self-explanatory: what might be meant by "True value of personal property" in the 1860 census? Students who are not themselves members of the Disciples of Christ denomination were almost certain to suppose that "Total number of Christian Churches" embraced all species of Christianity rather than a particular Protestant denomination.

The hypertext capability, to which much of the programming effort was devoted (section D.2.), allows us to address both these problems together. We created a topic hierarchy with five major headings: Economy, Society, Culture, Government and Environment. Under each major heading there are several subtopics; for example, under Society the three major subheadings are Stratification, Life Cycle and Ascription. Under Ascription we have data on Race, Ethnicity and Gender. In a few instances there are as many as five levels in the topic hierarchy. A given variable can be associated with more than one position in the topic hierarchy. Thus for example, variables in the 1980 census on the income of families with female householders can be classified under Society/Ascription/Gender, Society/Life Cycle/Domestic Group and Society/Stratification/Income. Moreover, we can associate text, ranging from a one-sentence note to an elaborate essay, with any topic in the hierarchy and with any variable.

Major efforts were mounted in both the summers of 1990 and 1991 and less intense efforts during the academic years, to do the requisite research and writing. Definitions and explanations have been written for most of the problematic variables, and a good deal of additional material about more general topics has also been generated. The students most heavily involved in this effort were Jared Day, a history graduate student, and Katherine Rashid, an undergraduate history major.

D.4. Data preparation

Our intention was to add U.S. census data for 1790 to 1830 and election data for 1840 to 1972 to the system and, with the assistance of Professor Gilbert Shapiro, to develop some data on French history since the eighteenth century. Although Professor Shapiro did accomplish some important spade work for an eventual French history application, higher priority was given to the U. S. data.

The most labor intensive part of the data preparation was hand-editing of the machine-readable maps. Prior to the FIPSE funding we had created a machine-readable map of the U.S. by counties which contained boundary changes at 10-year intervals since from 1840 to 1980. If a county was created between, say, 1870 and 1880, however, it was necessary to fix the date of the change to the nearest two years to accommodate the biennial election data. For the period prior to 1840 it was necessary to do substantial new digitizing with a drawing editor. These tasks were carried out under the leadership of Julie Smith, a history graduate student.

Retrieving the pre-1840 census data from archive tapes and formatting it presented no serious difficulites. However, the 1840-1972 election data were more problematic because they were stored in a single rectangular file containing some 756 variables for 3237 counties or county equivalents. The distributed, workstation-based computing environment in which we are accustomed to working proved unworkable for partitioning and formatting this large a file.

A mainframe solution seemed plausible, but at the time when we realized that we needed to seek another technical solution Roy Wilson, a student in our applied history graduate program who has a background in systems engineering, expressed an interest in working on a project which would enable him to use the facilities of the Pittsburgh Supercomputing Center (PSC). Accordingly we applied for

and received an allocation of time on the Connection Machine, one of PSC's supercomputers. Wilson adapted a program written by Miller to take advantage of the Connection Machine's architecture. After considerable debugging, the program ran successfully. However, the delays caused by our false start in a workstation environment meant that the formatted election data became available to us quite late in the funding period.

D.5. Overall project direction

FIPSE paid for one month of Miller's time during the funding period to coordinate the work. In fact Miller spent significantly more than one person-month on this effort between August 1989 and February 1992, but he spent a smaller percentage of his time on it in that period than he had during the preceding four years.* Between the summer of 1985 and the summer of 1989, Miller had devoted to this project virtually all the time which a CMU faculty member would ordinarily spend on research. During that period Miller himself sometimes even intervened to carry out specific low-level data management and programming tasks in order to achieve smoother work flow.

In an institution like CMU, four years is a very long time even for a senior faculty member to forego research activity in his discipline. By the time FIPSE funding was received (which was nearly two years after our first approach to FIPSE) Miller was under institutional pressure to resume an active research program in his field (the history of modern Ireland) and indeed those pressures coincided with his own inclinations. In fact, for the past two and a half years Miller been actively involved in research which, while it builds on the skills developed through the GAHM project, is unrelated to the specific objectives of that project.

The result is that while the specific tasks outlined above have been accomplished as indicated, sometimes the work flow has not been ideal. Miller made a decision early in the project that he should not step in and carry out low-level, labor-intensive technical and editorial tasks at the expense of his research program. This means that although the components of the project are essentially complete,

---

* Miller estimates that he was spending over 60% of his professional effort, including uncompensated time during summers, on GAHM-related activities during the years immediately prior to FIPSE funding.

the product is in what the toy industry might call a "some assembly required" condition. Had Miller been operating in the way he did four years ago, he would probably have spent several person-months (very cost-ineffectively) during the past year being sure all the components fit together. In fact, those tasks are now left to the Maryland group (with whom, of course, Miller will consult as needed). Indeed, negotiating the arrangement with the Maryland group was a far better use of Miller's time from FIPSE's standpoint than the alternative.

E. Project Results

As we indicated in section A, this project is perhaps unusual among FIPSE projects in that its audience is outside the institution where the work is being done. We had hoped, however, to be able to carry out an evaluation in two sections of our CMU course in which one section would use the pre-FIPSE version of the software, and the other would use the version with the hypertext capabilities. Various circumstances described in Section D meant that we simply did not have the new version in a stable form ready for student use in time to carry out such an evaluation in the fall, 1991, semester, the last term which fell within our extended project period.

During the past year a good deal of attention has been paid to dissemination. We turned down an approach which would have packaged GAHM as essentially freeware on a compact disk to be distributed by Apple for its own marketing purposes. We also had discussions with Intellimation about the possibility of their becoming the publisher. We concluded that they had insufficient understanding of our product and its potential and inadequate commitment to product maintenance. Accordingly, we decided to turn over the maintenance, upgrading and marketing of GAHM to the Academic Software Development Group, headed by Dr. Chad McDaniel, at the University of Maryland.

This group had already carried out a port of the pre-FIPSE version of GAHM to Windows. McDaniel clearly has both the entrepreneurial skills and an appropriate professional background (as a quantitative anthropologist) to assume direction of the project. His group has the infrastructure to carry out the necessary technical functions and experience in developing and disseminating postsecondary educational computing applications. McDaniel has excellent contacts in both the world of academic computing and in the relevant

academic disciplines. We trust that FIPSE will cooperate in whatever way possible to ensure the success of this dissemination initiative.

F. Summary and Conclusions

The project succeeded in accomplishing its main goal which was to enhance a software product on which a new form of teaching U.S. history has been built at CMU, in order to promote its dissemination to other postsecondary institutions. That success is attested to by the fact that the Academic Software Development Group at the University of Maryland has mounted a major effort to disseminate GAHM on both Windows and the Macintosh. We are grateful for support from FIPSE which enabled us to bring to a satisfactory conclusion, as far as CMU is concerned, an educational development effort of some seven years' duration.

Our "advice to other practicioners who are interested in your project" is to contact Chad McDaniel, Computer Science Center, Bldg 224, University of Maryland, College Park, MD 20742 (telephone 301-405-2920), for the latest information on the availability of our product on popular platforms.

G. Appendices

What lessons have been learned which might be relevant to other FIPSE projects? Certainly the first lesson is not to underestimate the programming resources needed for a software development project. If we were to do one thing differently it would be to budget initially for 100% coverage of the programmer's salary for the entire funding period. The recurrent need to seek other sources of income for the programmer led to discontinuities in his effort on the project which made it difficult to achieve project goals in an orderly fashion. Moreover, it made us unnecessarily vulnerable to the consequences of rapid changes in CMU's goals in the area of educational computing.

Educational software development, when undertaken by a subject-matter specialist rather than a computer scientist, has to compete with other claims upon that person's professional attention. University administrators sometimes take intiatives which seem to reorder those claims in favor of software development, but such reordering is quite impermanent. Not only is it likely to conflict with priorities set by academic units within the university and with peer pressures within the individual's discipline, but it is also vulnerable to rapid and unpredictable shifts in technology and in the university's larger strategies with respect to computing. To take maximum advantage of important initiatives in educational computing, FIPSE would do well to respond earlier rather than later in the development cycle of such projects to requests for funding of elements of those efforts which are consistent with the Fund's objectives.